Things for the Link Report

Alvin Sheng

Executive Summary

Questions of Interest

Statistical Analysis

Takeaways

Data Pre-processing

We only know the click count for each link address, not the actual link, within a newsletter. Thus, if there are multiple links with the same address in a newsletter, we don't know how many clicks each of those links received. We assume that the first text link with a given address received all the clicks associated with that address. We distinguish between unique links (those with addresses that appear only once) and duplicated links (those with addresses appearing multiple times) in our model.

Features used in Link Model

We created several features for text links:

1. Bolded: whether the image is bolded.

2. Font Size: ranges from 10-48 point.

3. Font Color: 38 possible colors.

4. Image Associated: there is an image within the newsletter with the same link address.

5.

6.

There is one feature for image links:

1. Image Width: width of an image on a computer browser, in pixels. This is an indicator of how large the image is. The image height depends on the width and the aspect ratio of the image.

Merging